Optimization of an Ill-posed Problem

M. Allmaras S. Indratno K. Kazmi

07/20/2007

MSRI Workshop on Continuous Optimization, 2007

Outline

Problem

Our goal is to recover the function

$$x_{\text{exact}}(t) = \begin{cases} 1, & \text{for } t \in \left[\frac{1}{4}, \frac{1}{2}\right) \\ 0, & \text{otherwise} \end{cases}$$

from a given set of measurements

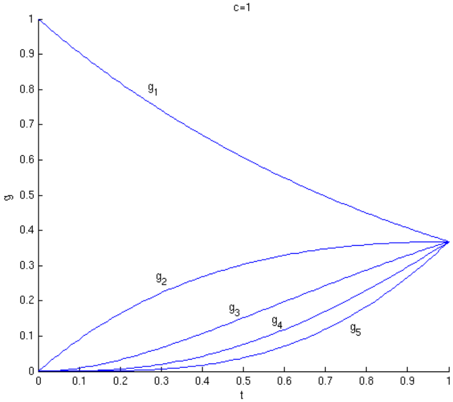
$$d_k = \int_0^1 g_k(t) x_{\text{exact}}(t)\, dt\,, \qquad k = 1, \ldots, m,$$

where

$$g_k(t) = t^{k-1} e^{-ct}\,, \qquad k = 1, \ldots, m$$

with a constant c.

Kernel functions



Formulation of the problem

In operator notation, we have

$$Gx = d \quad \text{with}$$
$$G : L^2[0,1] \rightarrow \mathrm{R}^m, \quad (Gx)_k = \langle g_k, x \rangle$$

Optimization problem

We look for the solution of $Gx = d$ with minimal norm:

$$\min \; N(x) \quad \text{subject to} \quad Gx = d$$

where

$$N(x) := \frac{1}{2} \|x\|^2$$

This is a quadratic optimization problem with linear equality constraints.

Optimality conditions

From the Lagrangian

$$L(x, \lambda) = \frac{1}{2} \|x\|^2 + \lambda^T (d - Gx)$$

we get the optimality conditions

$$\nabla_x L(x, \lambda) = 0 \quad \Leftrightarrow \quad x - G^* \lambda = 0$$

$$Gx = d$$

Optimality conditions

These conditions are satisfied by

$$GG^*\lambda = d$$

and

$$x = G^*\lambda = \sum_{k=1}^{m} \lambda_k g_k$$

where

$$GG^* : \mathrm{R}^m \to \mathrm{R}^m$$

Linear system

We have reduced the infinite dimensional optimization problem to solving a system of m linear equations:

$$GG^*\lambda = d\,.$$

Because the g_k are linearly independent, GG^* is invertible and we get the unique solution

$$x = G^+ d = G^*(GG^*)^{-1} d$$

Gram matrix

Because

$$(Gx)_i = \langle g_i, x \rangle \quad \text{and} \quad G^*\lambda = \sum_{j=1}^{m} \lambda_j g_j \ ,$$

GG^* is an $m \times m$ matrix with entries

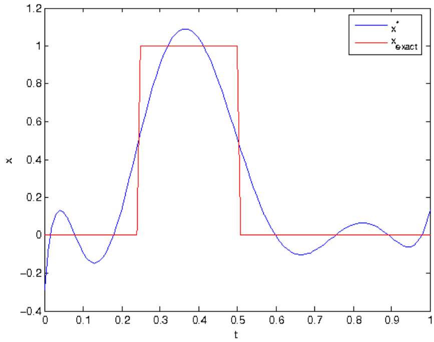
$$(GG^*)_{ij} \ = \ \langle g_i, g_j \rangle$$

These inner products can be evaluated as

$$\langle g_i, g_j \rangle = \frac{\Gamma(i+j-1)}{(2c)^{i+j-1}} \, \Gamma_{\text{inc}}(2c, i+j-1)$$

Solution

For $m = 9$ and $c = 1$ we get the following reconstruction x^*:



$$\|x_{\text{exact}}\|^2 = 0.25\,, \quad \|x^*\|^2 = 0.22$$

Measurement noise

In real experiments, the measurements d are usually contaminated with noise:

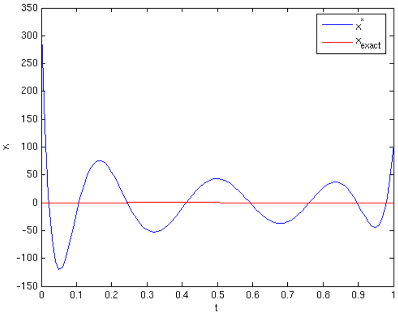
$$\hat{d}_k = \int_0^1 g_k(t) x_{\text{exact}}(t)\, dt + n_k\,, \qquad k = 1, \ldots, m,$$

where we assume n_k to be independent Gaussian noise with zero mean and standard deviation

$$\sigma_k = \frac{|d_k|}{10}.$$

Measurement noise

Using the noisy data $\hat{d}$ we get the following result:



$$\|x_{\text{exact}}\|^2 = 0.25\,, \quad \|x^*\|^2 = 2372.9$$

Ill-conditioning

The bad reconstruction in this case is due to the condition number of GG^*:

$$\kappa(GG^*) = 3.5176\text{e+}11$$

This ill-conditioning results from the fact that G is a compact integral operator and has unbounded inverse.

Regularization

Since the only available data is $\hat{d}$, the best we can do is to minimize the residue $\|Gx - \hat{d}\|$. Hence we try to solve the following optimization problem:

$$\min \frac{1}{2}\|x\|^2 + \frac{\mu}{2}\|Gx - \hat{d}\|_{S^{-1}}^2$$

where

$$S = \operatorname{diag}\left(\sigma_1^2, \ldots, \sigma_m^2\right)$$

is the covariance matrix of the noise n_k.
The regularization parameter μ is chosen to balance the norm of x and the datafitting accuracy.

Optimality conditions

The optimality condition for this unconstrained problem reads

$$x + \mu G^* S^{-1} G x - \mu G^* S^{-1} d = 0$$

One can show that

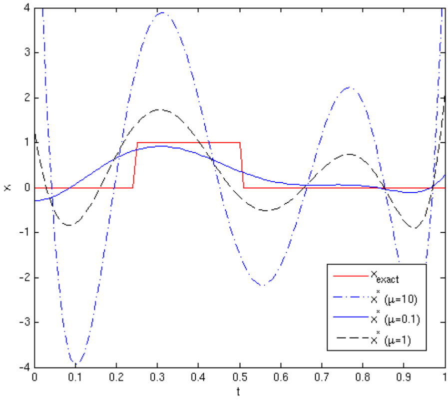
$$x = G^* \hat{\lambda} = \sum_{k=1}^{m} g_k \hat{\lambda}_k \,,$$

for some coefficient vector $\hat{\lambda} \in \mathrm{R}^m$, from which it follows that

$$\left(I + \mu S^{-1} G G^* \right) \hat{\lambda} = \mu S^{-1} d$$

Solution

Solution for various parameters μ:



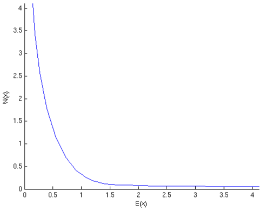
Influence of μ

To understand the influence of the regularization parameter μ on the solution, we can plot a tradeoff curve

$$\Big(E(x_\mu)\ ,\ N(x_\mu)\Big)$$

for varying μ-value, where

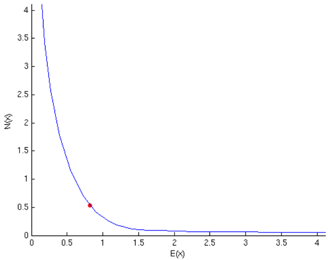
$$E(x) = \|Gx - \hat{d}\|^2_{S^{-1}},$$
$$N(x) = \|x\|^2.$$



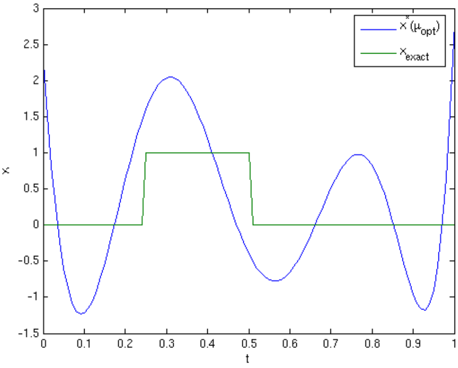
Optimal μ-parameter

By finding the point on the tradeoff curve nearest to the origin we can get an "optimal" parameter value μ^*. This amounts to solving the minimization problem

$$\min_{\mu \geq 0} \; E(x_\mu)^2 + N(x_\mu)^2$$



Solution for optimal parameter



$$\mu_{\text{opt}} = 1.5442$$